SAR Tomography at the Limit: Building Height Reconstruction Using Only 3–5 TanDEM-X Bistatic Interferograms

Yilei Shi, *Member, IEEE*, Richard Bamler, *Fellow, IEEE*, Yuanyuan Wang, *Member, IEEE*, and Xiao Xiang Zhu, *Senior Member, IEEE*

***Abstract*— Multibaseline interferometric synthetic aperture radar (InSAR) techniques are effective approaches for retrieving the 3-D information of urban areas. In order to obtain a plausible reconstruction, it is necessary to use more than 20 interferograms. Hence, these methods are commonly not appropriate for large-scale 3-D urban mapping using TanDEM-X data, where only a few acquisitions are available in average for each city. This article proposes a new SAR tomographic processing framework to work with those extremely small stacks, which integrates the nonlocal filtering into SAR tomography inversion. The applicability of the algorithm is demonstrated using a TanDEM-X multibaseline stack with five bistatic interferograms over the whole city of Munich, Germany. A systematic comparison of our result with TanDEM-X raw digital elevation models (DEMs) and airborne LiDAR data shows that the relative height accuracy of two-third buildings is within 2 m, which outperforms the TanDEM-X raw DEM. The promising performance of the proposed algorithm paved the first step toward high-quality large-scale 3-D urban mapping.**

***Index Terms*— 3-D urban mapping, digital elevation models (DEMs), interferometric synthetic aperture radar (InSAR), SAR tomography (TomoSAR), TanDEM-X.**

Manuscript received May 16, 2019; revised September 22, 2019, October 28, 2019, December 11, 2019, and February 22, 2020; accepted March 9, 2020. This work was supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme under Grant ERC-2016-StG-714087 (acronym: So2Sat, www.so2sat.eu), in part by the Helmholtz Association under the framework of the Young Investigators Group "SiPEO" under Grant VH-NG-1018 (www.sipeo.bgu.tum.de), in part by Munich Aerospace e.V. – Fakultät für Luft- und Raumfahrt, in part by the Bavaria California Technology Center (Project: Large-Scale Problems in Earth Observation), and in part by the Gauss Centre for Supercomputing (GCS) e.V. *(Corresponding author: Xiao Xiang Zhu.)*

Yilei Shi is with the Chair of Remote Sensing Technology, Technische Universität München (TUM), 80333 Munich, Germany (e-mail: yilei.shi@tum.de).

Richard Bamler is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany, and also with the Chair of Remote Sensing Technology, Technische Universität München (TUM), 80333 Munich, Germany (e-mail: richard.bamler@dlr.de).

Yuanyuan Wang is with Signal Processing in Earth Observation, Technische Universität München (TUM), 80333 Munich, Germany (e-mail: yuanyuan.wang@dlr.de).

Xiao Xiang Zhu is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany, and also with Signal Processing in Earth Observation, Technische Universität München (TUM), 80333 Munich, Germany (e-mail: xiaoxiang.zhu@dlr.de).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2020.2986052

I. Introduction

A. TanDEM-X Mission

TanDEM-X satellite is a German civil and commercial high-resolution synthetic aperture radar (SAR) satellite, which has almost identical configuration as its "sister" TerraSAR-X satellite. Together with TerraSAR-X, they are aiming to provide a global high-resolution digital elevation model (DEM) [1]. Both satellites use a spiral orbit constellation to fly in tight formation in order to acquire the image pair simultaneously, which significantly reduces the temporal decorrelation error and the atmospheric interference. Since its launch in 2010, TanDEM-X has been continuously providing high-quality bistatic interferograms that are nearly free from deformation, atmosphere, and temporal decorrelation.

B. SAR Tomography Techniques

Tomographic synthetic aperture radar (TomoSAR) is a cutting-edge SAR interferometric technique that is capable of reconstructing the 3-D information of scatterers and retrieving the elevation profile. Among the many multibaseline interferometric synthetic aperture radar (InSAR) techniques, TomoSAR is the only one that strictly reconstructs the full reflectivity along the third-dimension elevation. SAR tomography and its differential form (D-TomoSAR) have been extensively developed in the last two decades [2], [3] [4]–[9]. They are excellent approaches for reconstructing the urban area and monitoring the deformation, especially when using high-resolution data such as TerraSAR-X [10], [11] or COSMO-Skymed [12]. Compared to the classic multibaseline InSAR algorithms, compressive sensing (CS)-based methods [13], [14] can obtain extraordinary accuracy for TomoSAR reconstruction and show the super-resolution (SR) power, which is very important for urban areas, since layover is dominant.

Although TanDEM-X bistatic data have many advantages, there are only a limited number of acquisitions available for most areas. For a reliable reconstruction, SAR tomography usually requires fairly large interferometric stacks ($>$ 20 images) because the variance of the estimates is asymptotically related to the product of SNR and the number of acquisitions. Therefore, it is not appropriate for the microstacks, which have a limited number of interferograms [15].

C. Proposed Framework

As mention in Section III-C, the accuracy of 3-D reconstruction replies on the product of SNR and the number of measurements N. Since the motivation of this article is the large-scale urban mapping, the data we adopted are TanDEM-X strip map coregistered single-look slant-range complex (CoSSC), whose resolution is about 3.3 m in the azimuth direction and 1.8 m in the range direction. The typical number of available interferograms for most areas is 3–5 [16]. In [17], the pixels with similar height are grouped for the joint sparsity estimation, which leads to an accurate inversion of TomoSAR using only six interferograms. Although the unprecedented result is obtained, the accurate geometric information is usually not available for most areas. Therefore, the feasible way to keep the required precision of the estimates is to increase the SNR.

Recent works [18]–[20] show that SNR can be dramatically increased by applying nonlocal filters to the TomoSAR processing for different sensors, such as airborne E-SAR, COSMO-Skymed, and TerraSAR-X. In [18], different nonlocal filters have been adopted to improve the estimation of the covariance matrix for distributed scatterers, which leads to a better height estimation for simulated data and airborne SAR data. Ferraioli *et al.* [19] introduced the nonlocal filter and the total variation regularizer to improve the multibaseline phase unwrapping process. In [20], it is shown that we can achieve a reasonable reconstruction using only seven interferograms and better SR properties when the number of interferograms is relatively low.

In this article, we extend the concept of nonlocal CS TomoSAR in [20]–[22] and propose a new framework of spaceborne multibaseline SAR tomography with TanDEM-X bistatic microstacks, i.e., 3–5 interferograms. The framework includes nonlocal filtering, spectral estimation, model selection, and robust height estimation. Since different spectral estimators have a different estimation accuracy and computational cost, we compared the estimation accuracy of different estimators with microstacks.

The demonstration of different TomoSAR inversion methods for a large-scale area has shown in [11], [23], and [24]. Only a few works on the validation of single buildings were reported in [8], [10], and [25]. Therefore, the validation of the specified quality of the TomoSAR result at a larger scale would be of considerable interests for the scientific and commercial users. We choose Munich city as a test site because of a high-quality LiDAR reference available to us, and we propose a complete workflow to compare the TomoSAR point cloud [26] generated by the proposed framework, TanDEM-X DEM product, and LiDAR data.

D. Contribution of This Article

The major contributions of this article are summarized as follows.

1) We make possible a new application of bistatic SAR data for global building height reconstruction.
2) We have pointed out that for pixelwise multimaster TomoSAR, the well-known system equation is no longer valid in the multiscatterer case.
3) We have developed a framework for tomographic stacks with only 3–5 interferograms. A systematic investigation on the estimation accuracy and SR power for the microstacks has been carried out, which was never done before.
4) We use five TanDEM-X bistatic data to demonstrate the proposed framework. A systematic validation for a large-scale TomoSAR reconstruction has been carried out, and a method for comparing with other reference data is established. The results are quantitatively compared with the LiDAR reference for more than 34 000 buildings.

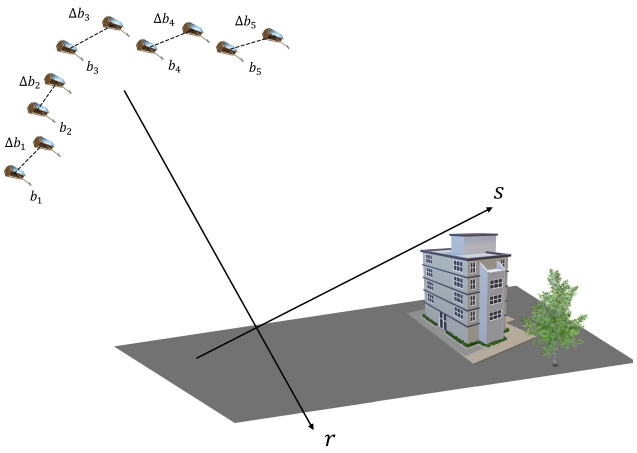


Fig. 1. Illustration of multimaster multibaseline SAR imaging.

This article is organized as follows. In Section II, the nonlocal TomoSAR framework is introduced. In Section III, the estimation accuracy of TomoSAR with small stacks has been systematically studied. The experiments using real data are presented in Section IV. In Section V, the quantitative validation is carried out. Finally, conclusions are given in Section V.

II. Nonlocal TomoSAR for Multimaster InSAR

In this section, we introduce the nonlocal TomoSAR framework for multimaster multibaseline InSAR configuration. Fig. 1 illustrates multimaster multibaseline SAR imaging. The framework consists of several steps: 1) nonlocal filtering; 2) spectral estimation; 3) model selection; and 4) robust height estimation. Fig. 2 shows the flowchart of the nonlocal TomoSAR framework.

A. Multimaster TomoSAR Imaging Model

For a fixed azimuth–range position, $\gamma(s)$ represents the reflectivity profile along elevation s. The measurement g_n, i.e., the complex value of the corresponding azimuth–range pixel, in the nth SAR image is then a sample of $\Gamma(k)$—the Fourier transform of $\gamma(s)$, where the elevation wavenumber k is a scaled version of the sensor's position b_n projected on the

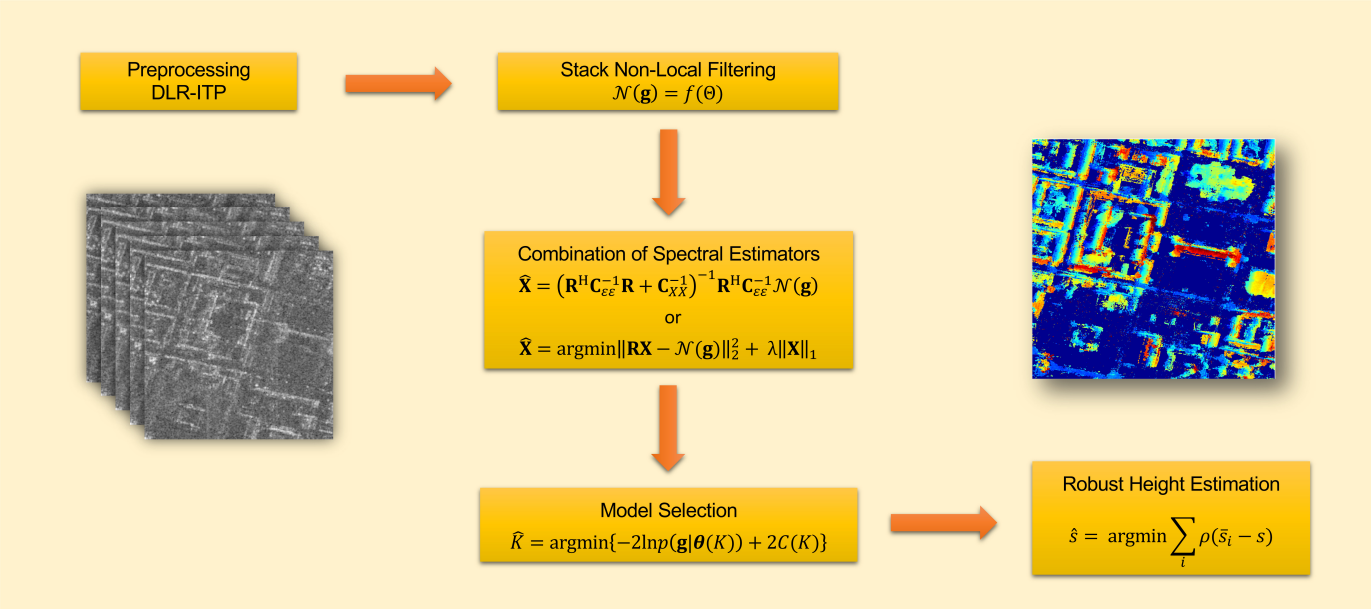


Fig. 2. Workflow of nonlocal TomoSAR framework.

cross-range–azimuth axis $b||s$

$$g_n = \Gamma(k_n) = \int \gamma(s)\exp(-jk_n s)ds \tag{1}$$

with

$$k_n = -\frac{4\pi b_n}{\lambda r}. \tag{2}$$

Note that b_n are no baselines but the positions of the sensor with respect to some origin. In the case of monostatic multitemporal data stacks, a *single* master g_0 is chosen with b_0, and its phase is subtracted from all other acquisitions: $g_n g_0^*/|g_0|$. This operation renders the phase spatially smooth and is a prerequisite for spatial phase unwrapping, averaging, and graph (network) processing. It does not reduce information since the phase of any (master) acquisition is random. Note that the choice of the point $b = 0$ only defines the x-r-s coordinate system. This point need not necessarily be the master track position b_0. However, $b_0 = 0$ is a mathematically convenient choice and is assumed in all conventional TomoSAR system model like in [27]. However, all the equations in [27] are actually independent of this particular choice. Only in the special case $b_0 = 0$, the b_n are identical to baselines.

Here, we are dealing with stacks of *bistatic* acquisitions, i.e., with the multimaster case. From each of these acquisitions, we get a master $g_{n,m} = \Gamma(k_n)$ taken at $b_{\text{master}} = b_n$ and a slave $g_{n,s} = \Gamma(k_n + \Delta k_n)$ image taken at $b_{\text{slave}} = b_n + \Delta b_n$, where Δb_n is the bistatic baseline (which takes the *effective* positions of the transmit–receive phase center into account). If we used a standard, i.e., single-master, TomoSAR inversion algorithm, we would confuse Δb_n and b_n. In the case of a single scatterer in $\gamma(s)$, this misinterpretation would do no harm because the Fourier transform of a single point has a constant magnitude and a linear phase. In order to determine the slope of the phase ramp, we can take any two samples and divide their phase difference by the difference in wavenumbers (= baseline). This is no longer true for two or more scatterers. The example of two symmetric and equally strong scatterers makes this clear

$$\begin{aligned}\gamma(s) &= \delta(s + s_0) + \delta(s - s_0)\\ &\updownarrow\\ \Gamma(k) &= 2\cos(s_0 k) = 2\cos\left(2\pi\frac{2s_0}{\lambda r}b\right).\end{aligned} \tag{3}$$

Hence, acquisitions with the same baseline Δb are different depending on where the two sensors were located along b. Every bistatic acquisition provides three pieces of information: the two magnitudes $|\Gamma(k_n)|$ and $|\Gamma(k_n + \Delta k_n)|$ as well as the phase difference $\angle\Gamma(k_n + \Delta k_n)\Gamma^*(k_n)$, i.e., we must normalize the phase by the respective master in every acquisition in order to become unaffected by deformation and atmospheric delay. However, spectral estimation-based conventional TomoSAR inversion algorithms require complex spectral samples at several wavenumbers, phase-normalized to a *single*-master phase. In a current parallel work by one of the authors [28], it is shown that pixelwise TomoSAR using multimaster acquisitions is a nonconvex hard to solve the problem.

This is true for *pixelwise* tomographic inversion or point scatterers. The situation becomes different although once we talk about averages of pixels, i.e., estimates of expectation values. Let us assume Gaussian-distributed scattering with a backscatter coefficient along elevation of

$$\sigma_0(s) = E\left\{|\gamma(s)|^2\right\}. \tag{4}$$

Assume that $\gamma(s)$ is white, its power spectral density is stationary and is the autocorrelation function of $\Gamma(k)$, i.e., the Fourier transform of $\sigma_0(s)$ as a function of the baseline wavenumber Δk

$$E\left\{\Gamma(k_n + \Delta k_n)\Gamma^*(k)\right\} = \int \sigma_0(s)\exp(-j\,\Delta k_n s)ds. \tag{5}$$

Instead of sampling the Fourier spectrum, we sample its autocorrelation function by the bistatic data stack. Since this

relationship is *independent* of $k \propto b$ because of stationarity, it makes no difference, where the two acquisitions have been taken, only their baseline Δb_n counts. In other words, we can use standard TomoSAR inversion algorithms in this case.

In this article, we use nonlocal filtering to improve SNR for microstacks. These filters perform ensemble averages with the number of looks on the order of tens to hundreds. Hence, we tend to the assumption that we work with reasonably good estimates of $E\{\Gamma(k_n + \Delta k_n)\Gamma^*(k)\}$ and can use the bistatic *interferograms* for TomoSAR reconstruction.

By introducing a noise $\boldsymbol{\varepsilon}$, the matrix notation of TomoSAR model can be formulated as

$$\mathbf{g} = \mathbf{R}\mathbf{X} + \boldsymbol{\varepsilon} \tag{6}$$

where $\mathbf{g} = [g_1, g_2, \ldots, g_n]^{\mathrm{T}}$ is the vector notation of the complex-valued measurement with dimension $N \times 1$, and $\mathbf{X} \sim \sigma_0(s_l) = E\{|\gamma(s_l)|^2\}$ is the expectation value of the reflectivity profile along elevation uniformly sampled at $s_l (l = 1, 2, \ldots, L)$. $\mathbf{R}$ is a sensing matrix with the dimension $N \times L$, where $R_{nl} = \exp(-j\,\Delta k_n s_l)$.

B. Nonlocal Procedure

Since we have only a limited number of acquisitions for a large-scale area, the SNR needs to be dramatically increased to obtain the required accuracy. As shown in [20], the nonlocal procedure is an efficient way to increase the SNR of interferograms without notable resolution distortion. The idea of patchwise nonlocal means considers all the pixels s in the search window; when the patch with the central pixel s is similar to the patch with the central pixel c, the value of s is selected for calculating the value of pixel c. The value of pixel c is estimated using a weighted maximum likelihood estimator (WMLE)

$$\hat{\boldsymbol{\Theta}}_c = \operatorname{argmax} \sum_s \mathbf{w}(i_s, j_s) \log p(\mathbf{g}_s|\boldsymbol{\Theta}) \tag{7}$$

where weights $\mathbf{w}(i_s, j_s)$ can be calculated by using patchwise similarity measurement [20]. Assume that we have two expressions $\mathbf{g} = (I_1, I_2, \phi)$ and $\boldsymbol{\Theta} = (\psi, \mu, \sigma^2)$, where $\mathbf{g}$ denotes the complex-valued measurement. I_1 and I_2 are the intensity of the two SAR images. ϕ is the interferometric phase. $\boldsymbol{\Theta}$ is the true value of the parameters, where ψ is the noise-free interferometric phase, μ is the coherence magnitude, and σ^2 is the variance. The likelihood function $p(\mathbf{g}_s|\boldsymbol{\Theta}) = p\big(I_{1,s}, I_{2,s}, \phi_s|\psi, \mu, \sigma^2\big)$ is adopted from [29] with the following formulation:

$$\begin{aligned} &p(I_1, I_2, \phi|\psi, \mu, \sigma^2) \\ &\quad = \frac{1}{16\pi^2\sigma^4(1-\mu^2)} \\ &\qquad \times \exp\left[-\frac{I_1 + I_2 - 2\sqrt{I_1 I_2}\mu\cos(\phi - \psi)}{2\sigma^2(1-\mu^2)}\right] \end{aligned} \tag{8}$$

where $\mathcal{N}(.)$ denotes the nonlocal estimator, where $\mathcal{N}(\mathbf{g}) = f(\hat{\boldsymbol{\Theta}})$. $\hat{\boldsymbol{\Theta}} = (\hat{\psi}, \hat{\mu}, \hat{\sigma^2})$ represents the parameters being estimated, where $\hat{\psi}$ is the estimated interferometric phase, $\hat{\mu}$ stands for the coherence magnitude, and $\hat{\sigma^2}$ stands for the variance. $f(\hat{\boldsymbol{\Theta}})$ is the maximum likelihood estimator, and the estimated parameters can be formulated as

$$\hat{\psi} = -\arg\left(\sum_s \mathbf{w}_s \mathbf{g}_{1,s} \mathbf{g}_{2,s}^*\right) \tag{9}$$

$$\hat{\mu} = \frac{2\sum_s \mathbf{w}_s |\mathbf{g}_{1,s}||\mathbf{g}_{2,s}|}{\sum_s \mathbf{w}_s \left(|\mathbf{g}_{1,s}|^2 + |\mathbf{g}_{2,s}|^2\right)} \tag{10}$$

$$\hat{\sigma^2} = \frac{\sum_s \mathbf{w}_s \left(|\mathbf{g}_{1,s}|^2 + |\mathbf{g}_{2,s}|^2\right)}{4\sum_s \mathbf{w}_s}. \tag{11}$$

The patch size and the search window size are set to be 7×7 and 21×21 according to the experimental study, which is also reported by other works [30], [31]. Each pixel represents 2.17 m in the azimuth direction and 1.36 m in the range direction.

C. Spectral Estimation

After the nonlocal procedure, spectral estimation is applied. The most relevant spectral estimation algorithms, including singular value decomposition (SVD) [5], [7] and CS, are introduced in the following.

1) *SVD:*

$$\hat{\mathbf{X}} = \left(\mathbf{R}^{\mathrm{H}}\mathbf{C}_{\varepsilon\varepsilon}^{-1}\mathbf{R} + \mathbf{C}_{XX}^{-1}\right)^{-1}\mathbf{R}^{\mathrm{H}}\mathbf{C}_{\varepsilon\varepsilon}^{-1}\mathcal{N}(\mathbf{g}). \tag{12}$$

2) *CS:*

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}}\left\{\|\mathbf{R}\mathbf{X} - \mathcal{N}(\mathbf{g})\|_2^2 + \lambda\|\mathbf{X}\|_1\right\} \tag{13}$$

where $\mathbf{C}_{\varepsilon\varepsilon}$ is the noise covariance matrix, which is defined as

$$\mathbf{C}_{\varepsilon\varepsilon} = (\mathbf{g} - \mathbf{R}\mathbf{X}) \cdot (\mathbf{g} - \mathbf{R}\mathbf{X})^{\mathrm{H}}. \tag{14}$$

Under the assumption that the model errors are circular Gaussian-distributed with zero mean, the noise covariance matrix is formulated as $\mathbf{C}_{\varepsilon\varepsilon} = |\sigma_\varepsilon|^2\mathbf{I}$ and $|\sigma_\varepsilon|^2$ is the noise power level. $\mathbf{C}_{XX}$ is the covariance matrix of the prior, if it is assumed to be white, i.e., $\mathbf{C}_{XX} = \mathbf{I}$.

The choice of different combinations of spectral estimators depends on the required accuracy, the computation time, and others. We follow the procedure proposed in [23]. It consists of three parts: 1) an efficient low-order spectral estimation; 2) the discrimination of the number of scatterers; and 3) an accurate high-order spectral estimation. The elevation profile is first estimated by an efficient low-order spectral estimator in order to discriminate the number of scatterers in one resolution cell. Then, CS-based approach is adopted for the pixel, which has multiple scatterers. This method decreases the amount of pixels that need the L_1-minimization, which leads to reduce the computational cost. Furthermore, the rest of the pixels can be efficiently solved by randomized blockwise proximal gradient method [24].

D. Model Selection

The abovementioned spectral estimators retrieve a nonparametric reflectivity profile. Since our data are in urban area, we assume only a few dominant scatterers exist along the reflectivity profile. Therefore, the number of scatterers $\hat{K}$ is

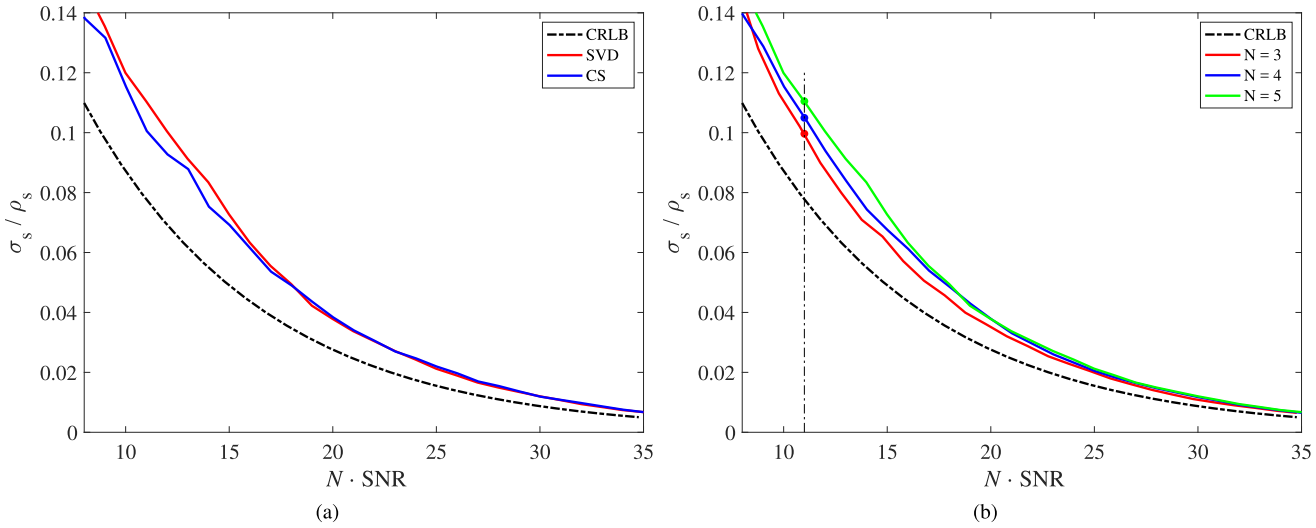


Fig. 3. Monte Carlo simulations of a single scatterer with SNR in [0 30] (dB). The X-axis presents $N \cdot$SNR in dB. The Y-axis is the normalized CRLB σ_s/ρ_s. (a) Comparison of CRLB with different spectral estimators with five acquisitions. SVD (red solid line), CS (blue solid line), and CLRB (black dashed-dotted line). (b) Comparison of CRLB using SVD with three–five acquisitions. $N = 3$ (red solid line), $N = 4$ (blue solid line), $N = 5$ (green solid line), and CLRB (black dashed-dotted line). The vertical black dashed-dotted line indicates the estimation accuracy for $N \cdot \text{SNR} = 11$ dB. The red, blue, and green markers represent $N = 3, 4, and 5$, respectively.

estimated by a model-order selection algorithm as well as their elevation in one azimuth–range pixel [7]. The estimator can be expressed as follows:

$$\hat{K} = \arg\min_{K}\{-2\ln p(\mathbf{g}|\boldsymbol{\theta}) + 2C(K)\} \tag{15}$$

where $C(k)$ is a model complexity penalty term, which avoids more complicated model overfitting the observed data. The classical penalized likelihood criteria are the Bayesian information criterion (BIC), the Akaike information criterion (AIC), and the minimum description length (MDL) principle [32].

As mentioned in [7], the criteria of the model-order selection have to be chosen according to the experiments for a particular situation because it is difficult to remove the bias of the selection.

E. Robust Height Estimation

To tackle the possible remaining outliers in the height estimates, the final height will be fused from the result of multiple neighboring pixels as a postprocessing. But instead of simple averaging, the height will be adjusted robustly using an *M-estimator*. Instead of minimizing the sum of squared residuals in averaging, M-estimator minimizes the sum of a customized function $\rho(.)$ of the residuals

$$\tilde{s} = \arg\min_{s} \sum_{i} \rho(\hat{s}_i - s) \tag{16}$$

where $\hat{s}_i$ is the elevation estimates of the ith neighboring pixel. It is shown that the closed-form solution of (16) is simply a weighted averaging of the heights of the neighboring pixels [33]. The weighting function can be expressed as follows: if the derivative of $\rho(x)$ exists:

$$w(x) = \frac{\partial\rho(x)}{x\partial x}. \tag{17}$$

The robust estimated height can be written as follows:

$$\tilde{h} = \frac{\sum_i w(x_i)\cdot\hat{h}}{\sum_i w(x_i)} \tag{18}$$

where $\hat{h} = \hat{s}\cdot\sin\theta$, and θ is the incident angle. The choice of the weighting function depends on the distribution of the heights. Without prior knowledge of the distribution, promising robust weighting functions are Tukey's biweight or t-distributed weighting [33]. For instance, the formulation of Tukey's biweight loss function can be written as follows:

$$\rho(x) = \begin{cases} -\dfrac{\left(c_r^2 - x^2\right)^3}{6c_r^4} + \dfrac{c_r^2}{6}, & |x| < c_r \\ \dfrac{c_r^2}{6}, & \text{elsewhere} \end{cases} \tag{19}$$

and the weighting function can be formulated as

$$w(x) = \begin{cases} 1 - \dfrac{x^4}{c_r^4} - \dfrac{2x^2}{c_r^2}, & |x| < c_r \\ 0, & \text{elsewhere.} \end{cases} \tag{20}$$

III. Estimation Accuracy of TomoSAR With Small Stacks

This section will discuss the theoretical 3-D reconstruction accuracy of a microstack with 3–5 interferograms. The estimation accuracy of TomoSAR has been systematically investigated. It is exhaustively shown in [15] that the elevation estimation accuracy and SR power depend asymptotically on the multiplication $N \cdot$ SNR. In this section, we investigate the estimation accuracy of TomoSAR with the extremely small number of interferograms, which is 3–5.

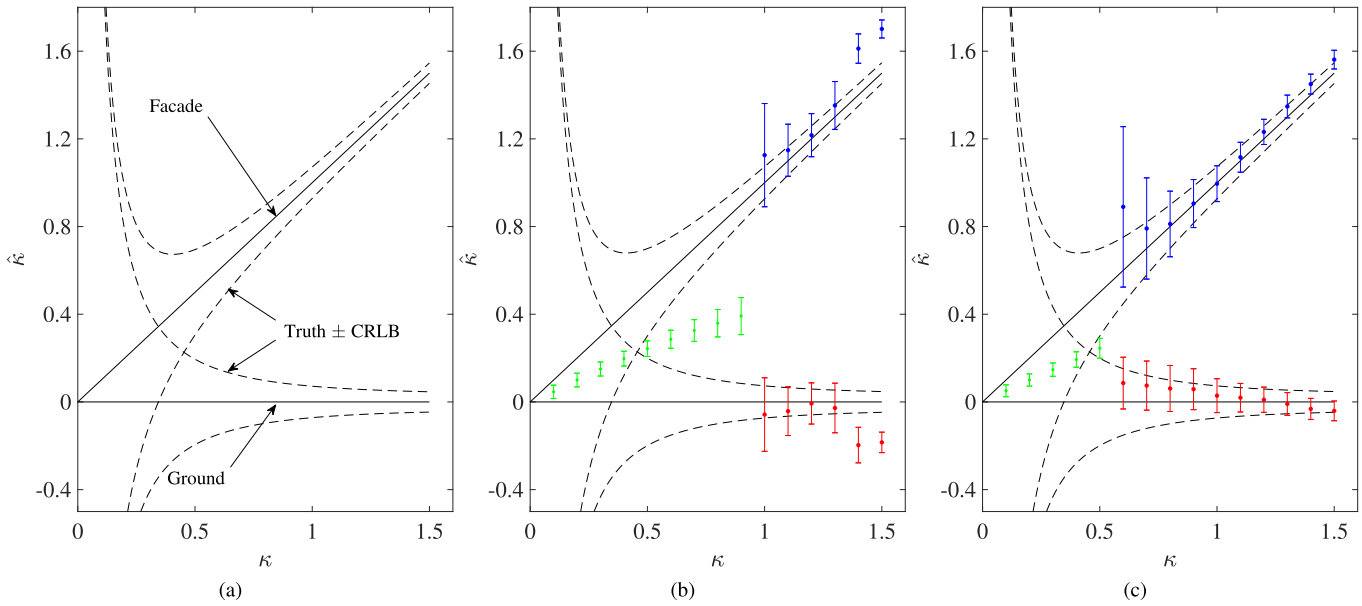


Fig. 4. Monte Carlo simulations of double scatterers with different normalized distances: $\kappa \in [0.1, 1.5]$ and SNR = 10 dB. The X-axis represents the normalized true distance κ of simulated facade and ground. The Y-axis is the normalized estimated distance $\hat{\kappa}$ of simulated facade and ground. The blue dot marker denotes the estimated location of facade, and the error bar indicates the standard deviation of the estimates, whereas the red dot marker represents the estimated location of ground. The green dot suggests that the detection rate of double scatterers is below 5% and denotes the estimated result of a single scatterer. (a) Illustration. (b) SVD. (c) CS.

A. Lower Bound for Microstacks

In the case of pixelwise TomoSAR inversion, i.e., without spatial averaging, each of our N bistatic pairs contains three pieces of information, as mentioned above. If we want to reconstruct elevation profiles containing M discrete scatterers, we need to infer $3M$ parameters, i.e., elevation, magnitude, and phase for each scatterer. Hence, an absolute lower bound of the microstack size is $N \geq M$.

On the other hand, distributed scatterers are characterized by only two parameters each: elevation and backscatter coefficient. Likewise, each interferogram provides only two parameters, magnitude and phase (difference). Since our goal is 3-D reconstruction based on bistatic data, we disregard motion-induced phase here. Hence, also in this case, the absolute lower limit is $N \geq M$. This limit is only a necessary condition, however, not sufficient from the robustness point of view, because of ambiguities in the inversion cost functions.

For 3-D urban mapping, the single- and double-scattering cases are the dominant ones. We investigate the cases $N = 3-5$ in this article because these are close to the mentioned limits and are relevant for TanDEM-X.

B. Cramer–Rao Lower Bound (CRLB)

It is demonstrated in [7] that the CRLB of the elevation estimates for single scatterer can be expressed as follows:

$$\sigma_s = \frac{\lambda r}{4\pi \cdot \sigma_b \cdot \sqrt{2 \cdot \mathrm{SNR} \cdot N}} \tag{21}$$

where σ_b is the standard deviation of the baseline distribution. N is the number of interferograms, and SNR is the signal-to-noise ratio.

For the double-scatterer case, the CRLB can be written as follows:

$$\sigma_{s_q} = c_0 \cdot \sigma_{s_q,0} \tag{22}$$

where $\sigma_{s_q,0}$ represents the CRLB on the elevation estimation of the qth scatterer without the interference with the others. c_0 is the correction factor of the interference for the scatterers, which are closely located [15]. It is nearly free from N and SNR, which can be written as follows:

$$c_0 = \max\left\{\sqrt{\frac{40\kappa^{-2}(1-\kappa/3)}{9-6(3-2\kappa)\cos(2\Delta\varphi)+(3-2\kappa)^2}}, 1\right\} \tag{23}$$

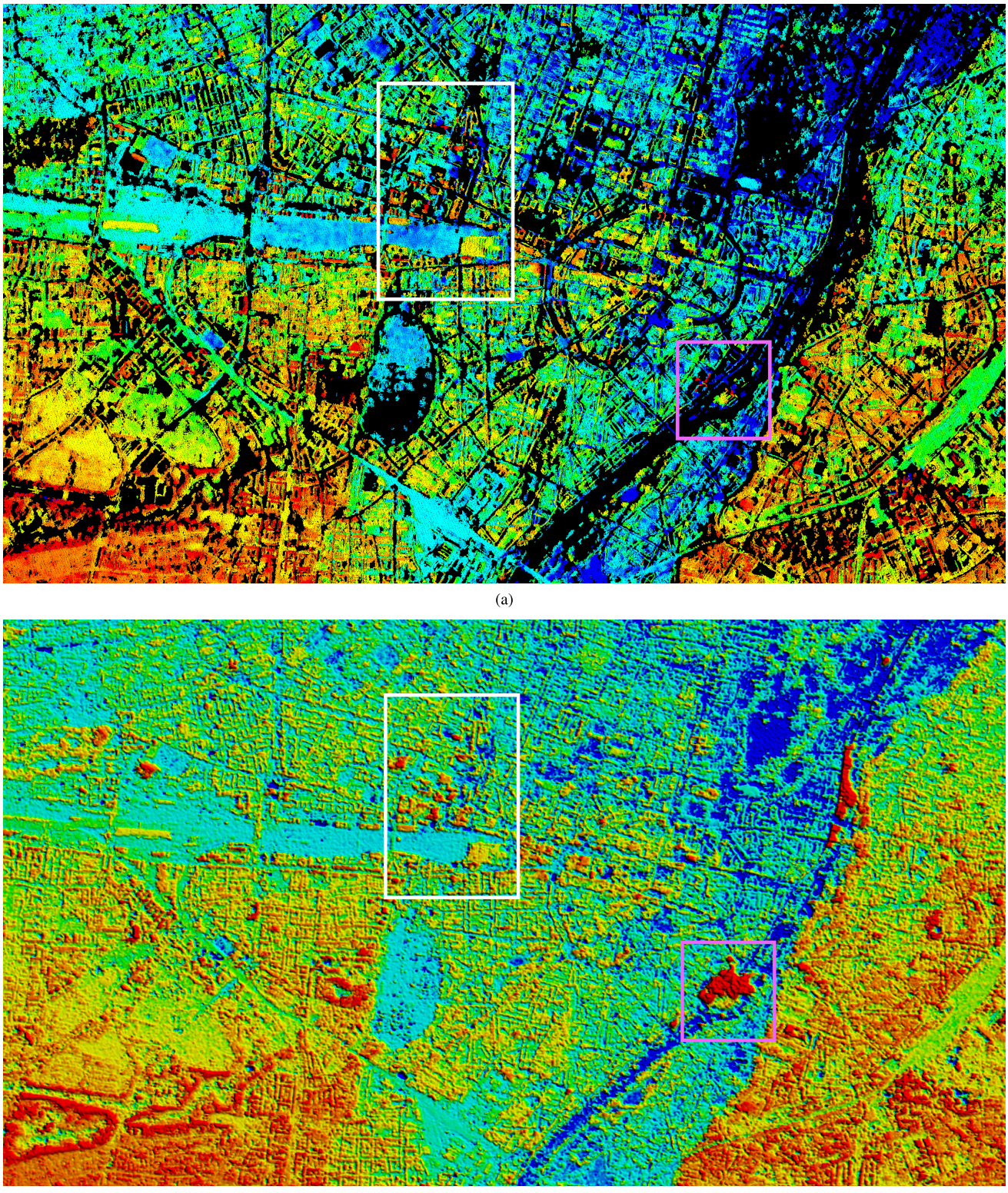
where $\Delta\varphi$ is the phase difference of the two scatterers. κ is the normalized distance between two scatterers (defined in the next section). Since $\Delta\varphi$ is a random variable, the approximated formulation of c_0 can be calculated by integrating the variances over $\Delta\varphi$

$$c_0 = \max\left\{2.57(\kappa^{-1.5}-0.11)^2+0.62, 1\right\}. \tag{24}$$

A note on the baseline distribution is worth mentioning: (21)–(24) are satisfied with large stacks. In the microstacks with only 3–5 acquisitions, the baseline distribution may be unfavorable, even if the baseline spread σ_b is acceptable. For example, if two baselines were very similar, the information content would be reduced. It is, therefore, desirable to have the baselines possibly statistically uniformly distributed.

C. Monte Carlo Simulations

In this section, we compare different spectral estimators using simulated data. Two cases were carried out. The first



(a)

(b)

Fig. 5. Visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM over Munich, Germany. Color code: 565 m (blue)–596 m (red), scene size: 15 km × 9 km, north = top. The violet bounding box indicates the region of interest (ROI) over the area of European bureau of patent, and the white bounding box indicates the ROI near Munich central station. (a) Point clouds generated by NL-TomoSAR with five interferograms. (b) TanDEM-X DEM.

case considers only a single scatterer in the interest of exploring the effect of N and SNR on the estimation accuracy for microstacks and the performance of different estimators. The second case considers double scatterers to investigate the estimation accuracy and the SR power for different estimators. The inherent (Rayleigh) elevation resolution ρ_s is inversely

TABLE I
PARAMETERS OF TANDEM-X STRIPE MAP ACQUISITION OF MUNICH

Name	Symbol	Value
Distance from the scene center	r	698 km
Wavelength	λ	3.1 cm
Incidence angle at scene center	θ	50.4°
Maximal elevation aperture	Δb	187.18 m
Number of interferograms	N	5

TABLE II
DETAILED INFORMATION OF TANDEM-X STRIPE MAP ACQUISITION FOR USED DATA SET

No.	Date	Baseline [m]	Height Ambiguity [m/cycle]
1	2016-07-25	184.40	50.30
2	2016-09-07	171.92	54.01
3	2017-02-19	32.30	286.03
4	2017-04-26	-2.78	-8710.99
5	2017-07-01	9.30	1073.03

proportional to the maximal elevation aperture Δb [15]

$$\rho_s = \frac{\lambda r}{2\Delta b}. \tag{25}$$

The normalized distance is defined as follows:

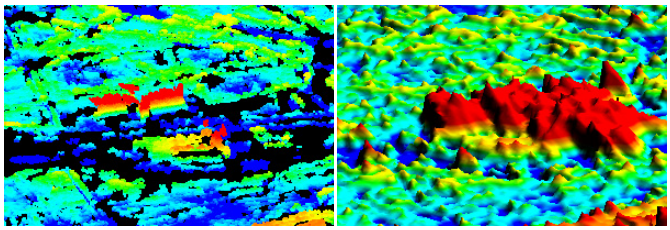
$$\kappa = \frac{s}{\rho_s}. \tag{26}$$

For the first test case, only one scatterer is placed at $s = 0$, and the SNR is in the range between 0 and 30 dB. For each $N \cdot \mathrm{SNR}$ value, 100 different baseline distributions were generated. We carried out a Monte Carlo simulation for each baseline distribution with 10 000 realizations. Afterward, the CRLB was evaluated by averaging the value of 100 different baselines. Fig. 3(a) shows a performance comparison between SVD and CS on simulated data with five acquisitions for a single scatterer. The X-axis presents $N \cdot \mathrm{SNR}$ in dB. The Y-axis is the normalized CRLB σ_s/ρ_s. As one can see, both approaches have similar estimation accuracy. They are asymptotically toward the CRLB and collapse it when $N \cdot \mathrm{SNR}$ is large. More interesting is when $N \cdot \mathrm{SNR}$ is fixed, leaving N as a variable. Fig. 3(b) presents the estimation accuracy of SVD with $N = 3, 4, 5$. It shows that the accuracy when $N = 3$ is the smallest. This indicates that SNR carries more weight than N on the estimation accuracy when N is very small.

With the 3-D reconstruction accuracy of a single scatterer clearly analyzed, we switch to the double-scatterer case. In the simulation, the elevation of one scatterer is fixed at 0. The normalized elevation of the other scatterer is increased from 0.1 to 1.5 to mimic the layover of a ground layer and a facade layer. The number of acquisitions is set to $N = 3$–5 as same as the first simulation. SNR is set to be 10 dB since the SNR of TanDEM-X bistatic data is usually higher than this value in urban area [34].

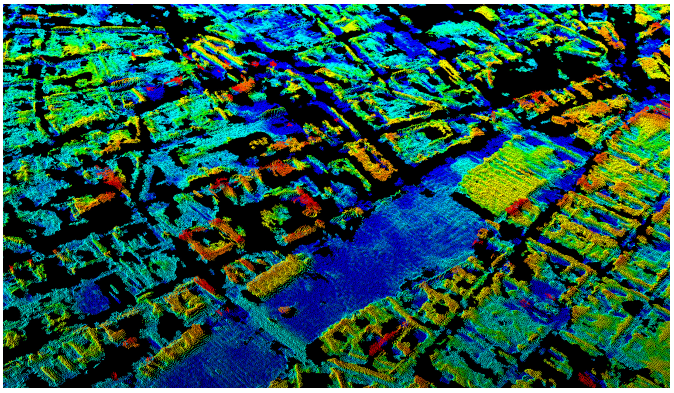
The Monte Carlo simulation result is shown in Fig. 4. The x-axis represents the true normalized elevation distance κ of the simulated facade and ground layers. The Y-axis is the estimated normalized elevation distance $\hat{\kappa}$ of the simulated facade and ground layers. The two solid lines in Fig. 4(a)–(c) represent the true position of the building facade and the ground, respectively. The dashed lines imply the true position plus and minus the CRLB. The blue bar and dot imply the standard deviation and the mean of the estimated elevation of the facade scatterers, whereas the red ones represent those of the ground scatterers. The green dot indicates that the detection rate of double scatterers is below 5% and denotes the estimated result of the single scatterer. Fig. 4(b) and (c) shows the estimated results by SVD and CS, respectively.

As one can see in Fig. 4(b) and (c), the result of SVD has a larger bias and slightly bigger standard deviation than CS. Note that, compared to SVD, CS can give a better result, not only the accuracy of the estimation but also the SR power. As one can see that the SVD has scarcely no SR power, which can only distinguish double scatterers tile one Rayleigh resolution ρ_s. In contrast, CS can achieve until 0.6 ρ_s.

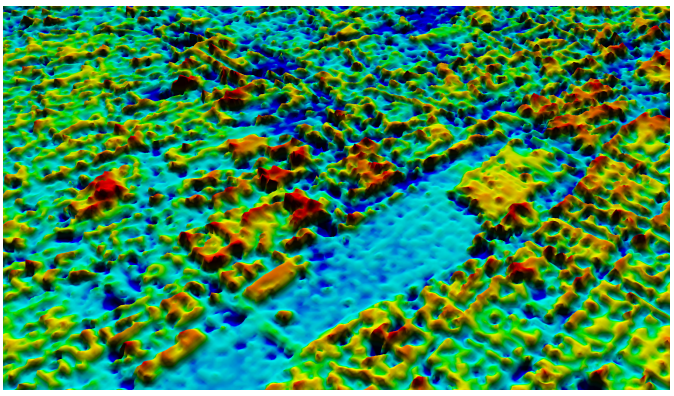


(a) (b)

Fig. 6. Visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM, close-up 3-D view over the area of European bureau of patent. (a) TomoSAR point clouds. (b) TanDEM-X DEM.



(a)



(b)

Fig. 7. Visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM, close-up 3-D view over the area of Munich central station. (a) TomoSAR point clouds. (b) TanDEM-X DEM.

Fig. 8. Optical images of the nine test sites for quantitative comparison of NL-TomoSAR point clouds and TanDEM-X DEM. (a) Munich central station. (b) European bureau of patent. (c) Technical University of Munich. (d) Railway signal light stand. (e) Train repair garage. (f) Residential building between two bridges. (g) Munich University of Applied Sciences. (h) Residential building near Lowenbrau beer company. (i) Karstadt (shopping mall).

IV. Practical Demonstration

A. Data Description

We make use of a stack of five coregistered TanDEM-X bistatic interferograms to evaluate the proposed algorithm. The data set is over Munich, Germany, whose slant range resolution is 1.8 m and the azimuth resolution is 3.3 m. The images were acquired from July 2016 to April 2017. The most pertinent parameters of a TanDEM-X bistatic stripe map acquisition of Munich are listed in Tables I and II. All the preprocessing steps, like deramping, are standard that are known from bistatic forest tomography. For interested readers, please refer to [2].

B. Visual Comparison With TanDEM-X Raw DEM

In this article, the TanDEM-X raw DEM is adopted for visual comparison with TomoSAR point clouds of the test area, which is formed by two TanDEM-X bistatic acquisitions using the integrated TanDEM-X processor (ITP).

A top view of the reconstructed point cloud of TomoSAR is shown in Fig. 5(a). The black regions in the figure are where the pixels are not coherent. The corresponding area of TanDEM-X raw DEM is presented in Fig. 5(b) as a comparison. It is clear that the result of TomoSAR point cloud preserves more detailed building structures. The road layer is also better represented in the TomoSAR result as well. In Fig. 5(b), the flat ground surface are well-reconstructed. But when it comes to complex or high-rise buildings, their accuracy is compromised. For instance, the building of European bureau of the patent in the bottom right (red) along the Isar river. A closed view of this building can be seen in Fig. 6. Due to the complex building structure, as well as the multilooking processing, the TanDEM-X raw DEM merges several buildings together and exhibits lower accuracy on the height of the buildings.

As another example, Fig. 7 shows the visual comparison over the area around Munich central station. It is clear that NL-TomoSAR result can show more detailed structures, such as the bridge, the central station, and roads.

V. Quantitative Validation

In this section, we have quantitatively compared the TomoSAR point clouds with TanDEM-X raw DEM, as well as a much more precise LiDAR reference. The LiDAR data set of Munich is provided by Bavarian State Office for Survey and Geoinformation with 10-cm accuracy [35].

Since the TomoSAR point cloud is with respect to a reference point that was chosen during the TomoSAR processing, its location is not with respect to a geocoordinate system. We coregistrated the point cloud of TomoSAR with the DEM and the LiDAR point cloud. In addition, in order to compare point clouds with DEM, we rasterize the two point clouds. These preparing steps are briefly explained in this section.

TABLE III

STATISTICS OF QUANTITATIVE COMPARISON OF THE NINE TEST STRUCTURES. FIRST COLUMN SHOWS THE NUMBER OF EACH STRUCTURE. SECOND COLUMN IMPLIES THE SOURCE OF EACH RESULT, I.E., t (TOMOSAR), l (LIDAR), AND d (DEM). THIRD AND FOURTH COLUMNS PRESENT THE STATISTICS (MIN, MAX, MEAN, AND STANDARD DEVIATION) OF SAMPLE POINTS AT TOP AND BOTTOM LAYERS. FIFTH COLUMN DEMONSTRATES THE RELATIVE HEIGHT OF EACH STRUCTURE, WHICH IS CALCULATED BY USING THE MEAN VALUE OF THE TOP LAYER MINUS THE MEAN VALUE OF THE BOTTOM LAYER. SIXTH COLUMN SHOWS THE RELATIVE HEIGHT DIFFERENCE BETWEEN TOMOSAR POINT CLOUDS AND LIDAR DATA, AS WELL AS BETWEEN TANDEM-X RAW DEM AND LIDAR DATA

Structures	Sources	Top				Bottom				Height	Absolute Height Difference
		Min	Max	Std	Mean	Min	Max	Std	Mean		
Structure 1	T	-3.76	2.59	1.22	-0.75	-19.30	-12.87	1.15	-15.26	14.51	**0.69**
	L	-	-	-	539.01	-	-	-	525.19	13.82	**-**
	D	583.19	597.58	2.32	587.91	566.16	570.15	2.01	568.06	19.84	**6.02**
Structure 2	T	20.39	22.62	0.56	21.39	-27.85	-22.62	1.18	-25.30	46.70	**0.75**
	L	-	-	-	559.09	-	-	-	513.14	45.95	**-**
	D	598.57	642.43	8.35	624.99	556.61	583.46	4.45	574.83	50.16	**4.21**
Structure 3	T	13.84	17.04	0.97	15.32	-25.52	-21.56	1.09	-23.67	38.49	**0.90**
	L	-	-	-	552.97	-	-	-	515.38	37.59	**-**
	D	594.31	599.13	2.12	596.15	562.03	569.28	1.60	565.18	30.97	**6.62**
Structure 4	T	-4.61	-1.90	0.74	-2.91	-13.84	-10.35	0.83	-12.05	9.14	**0.67**
	L	-	-	-	535.04	-	-	-	526.57	8.47	**-**
	D	582.41	584.94	0.84	584.06	572.76	574.96	0.48	573.42	10.64	**2.17**
Structure 5	T	-3.95	-0.96	0.63	-2.44	-15.94	-13.67	0.54	-14.61	12.17	**0.96**
	L	-	-	-	535.62	-	-	-	524.41	11.21	**-**
	D	583.26	587.26	0.96	584.77	572.71	578.98	1.22	575.58	9.19	**2.02**
Structure 6	T	10.42	13.47	0.49	12.11	-16.05	-14.31	0.82	-15.61	27.72	**0.67**
	L	-	-	-	551.73	-	-	-	523.34	28.39	**-**
	D	587.23	594.29	2.12	589.76	567.22	570.82	1.32	569.36	20.4	**7.99**
Structure 7	T	2.71	6.65	0.87	4.99	-27.57	-21.32	1.41	-24.25	29.24	**0.60**
	L	-	-	-	548.01	-	-	-	519.37	28.64	**-**
	D	574.71	597.30	5.21	588.27	563.98	571.78	2.26	568.09	20.18	**8.46**
Structure 8	T	0.06	6.42	1.34	4.06	-20.96	-20.27	0.18	-20.69	24.75	**0.94**
	L	-	-	-	542.96	-	-	-	517.27	25.69	**-**
	D	584.73	598.09	3.36	590.40	566.26	574.70	2.62	570.12	20.28	**5.41**
Structure 9	T	-7.53	-6.73	0.16	-7.41	-23.13	-22.57	0.29	-22.85	15.44	**0.89**
	L	-	-	-	530.39	-	-	-	515.84	14.55	**-**
	D	580.67	581.22	0.11	580.97	567.14	573.42	1.56	569.3	11.67	**2.88**

A. Geocoding

Since the result of TomoSAR inversion is a 3-D point cloud in the range–azimuth coordinate, the first step is to transform the result to Universal Transverse Mercator (UTM) coordinate with the range-Doppler approach [36].

B. Coregistration of Different Point Clouds

Consequently, when the TomoSAR point cloud is transformed to a UTM coordinate, its position may differ from the ground truth since the height of the reference point is unknown. Hence, the alignment of different point clouds is necessary. The most popular 3-D point cloud registration algorithm is iterative closest point (ICP) approach [37].

The performance of ICP depends on the initial alignment. Hence, a coarse alignment is adopted before applying ICP, which includes three steps: 1) the edge image is extracted by an edge detector, such as Sobel algorithm [38]; 2) the horizontal coregistration of two edge images is using cross correlation of two edge images; and 3) the vertical coregistration is using cross correlation of the two height histograms. After the coarse alignment, ICP can be applied for the fine alignment [39].

C. Object-Based Raster Data Generation

The direct comparison of TomoSAR and LiDAR points is not feasible, as the central position of two corresponding points (one TomoSAR and one LiDAR point) and the footprint of the points are differing. Consequently, for comparing both data, an object-based raster needs to be generated by using geographic information system GIS data.

D. Comparison of Individual Structure

In order to evaluate the estimation accuracy, nine test sites with high-average SNR have been chosen for individual quantitative comparison. Fig. 8 shows the optical images of nine test sites for quantitative comparison of NL-TomoSAR point clouds and TanDEM-X DEM. They are: 1) Munich central station; 2) European bureau of patent; 3) Technical University of Munich; 4) a railway signal light stand near Hirschgarten; 5) a train repair garage near Hirschgarten; 6) a residential building between two bridges (Hackerbrücke and Donnersbergebrücke); 7) Munich University of Applied Sciences; 8) a residential building near Lowenbrau beer company; and 9) Karstadt (shopping mall). The summary of the results is shown in Table III.

TABLE IV
STATISTICS OF QUANTITATIVE COMPARISON OF THE WHOLE CITY

Percentage of buildings	Estimation accuracy
38.7%	within 1 m
62.8%	within 2 m
93.3%	within 15 m

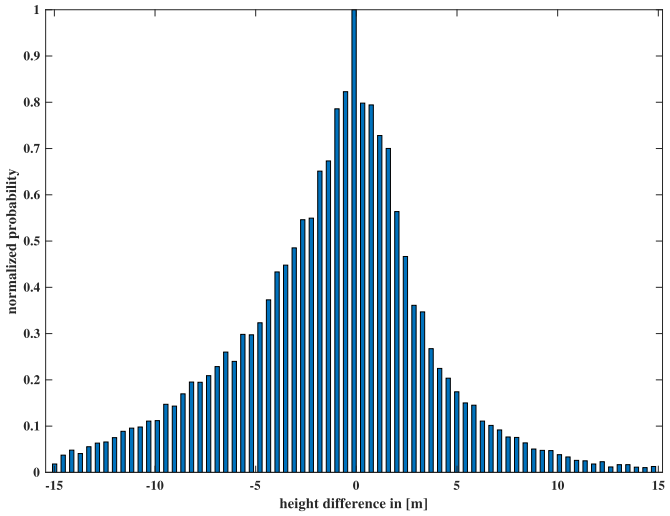


Fig. 9. Histogram of height differences of structures in the whole Munich area.

From Table III, we can see that the height differences between TomoSAR result and LiDAR data are within 1 m, and the height differences between TanDEM-X DEM product and LiDAR data vary from 2.5 to 8.5 m. Similar performance is shown in the standard deviation; for NL-TomoSAR, it is up to 1.4 m, and for TanDEM-X DEM, it is up to 8.4 m.

E. Average Accuracy

In order to have an assessment of the overall accuracy in a city scale, we compared all the 36 499 buildings in the area with the LiDAR point cloud. The 38.7% buildings are within 1-m accuracy. The 62.8% buildings are within 2-m accuracy. A detailed distribution of accuracy is listed in Table IV. However, the two data sets (TanDEM-X CoSSC and LiDAR) were acquired at different times. It is almost certain that changes happened during the period. Therefore, in order to obtain a more realistic assessment, we truncated the distribution of height difference at ± 15 m. The truncated histogram can be seen in Fig. 9. The 34 054 buildings remain after the truncation. Their overall standard deviation is 1.96 m.

VI. CONCLUSION

A new SAR tomographic inversion framework tailored for a very limited number of measurements is proposed in this article. A systematic investigation of the estimation accuracy of TomoSAR with microstacks is carried out using simulated data. Our experiments show that SVD and CS-based methods have almost identical performance on the estimation of single scatterer and the SNR plays a more important role than N for the estimation accuracy, when N is small. For the estimation of double scatterers, CS-based approach outperforms the other spectral estimators. Experiments using TanDEM-X bistatic data showed the relative height accuracy of 2 m can be achieved in a large scale. Thus, it demonstrates the proposed framework being a promising solution for high-quality large-scale 3-D urban mapping.

REFERENCES

[1] G. Krieger *et al.*, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3317–3341, Nov. 2007.
[2] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.
[3] F. Gini, F. Lombardini, and M. Montanari, "Layover solution in multibaseline SAR interferometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1344–1356, Oct. 2002.
[4] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.
[5] G. Fornaro, F. Lombardini, and F. Serafino, "Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 702–714, Apr. 2005.
[6] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.
[7] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.
[8] N. Ge, F. R. Gonzalez, Y. Wang, Y. Shi, and X. X. Zhu, "Spaceborne staring spotlight SAR Tomography—A first demonstration with TerraSAR-X," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3743–3756, Oct. 2018.
[9] X. Zhu, Y. Wang, S. Montazeri, and N. Ge, "A review of ten-year advances of multi-baseline SAR interferometry using TerraSAR-X data," *Remote Sens.*, vol. 10, no. 9, p. 1374, 2018.
[10] X. X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3150–3157, Aug. 2012.
[11] X. Xiang Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system," in *Proc. Joint Urban Remote Sens. Event*, Apr. 2013, pp. 159–162.
[12] G. Fornaro, A. Pauciullo, D. Reale, and S. Verde, "Multilook SAR tomography for 3-D reconstruction and monitoring of single structures applied to COSMO-SKYMED data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 2776–2785, Jul. 2014.
[13] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by L_1-norm regularization—The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.
[14] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-dimensional SAR focusing from multipass signals using compressive sampling," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 488–499, Jan. 2011.
[15] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, Jan. 2012.
[16] P. Rizzoli *et al.*, "Generation and performance assessment of the global TanDEM-X digital elevation model," *ISPRS J. Photogramm. Remote Sens.*, vol. 132, pp. 119–139, Oct. 2017.
[17] X. X. Zhu, N. Ge, and M. Shahzad, "Joint sparsity in SAR tomography for urban mapping," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1498–1509, Dec. 2015.
[18] O. D'Hondt, C. Lopez-Martinez, S. Guillaso, and O. Hellwich, "Nonlocal filtering applied to 3-D reconstruction of tomographic SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 272–285, Jan. 2018.
[19] G. Ferraioli, C.-A. Deledalle, L. Denis, and F. Tupin, "Parisar: patch-based estimation and regularized inversion for multibaseline SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1626–1636, Mar. 2018.
[20] Y. Shi, X. X. Zhu, and R. Bamler, "Nonlocal compressive sensing-based SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 3015–3024, May 2019.
[21] Y. Shi, X. X. Zhu, and R. Bamler, "SAR tomography using non-local sparse reconstruction," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 6091–6094.

[22] Y. Shi, Y. Wang, X. X. Zhu, and R. Bamler, "Non-local SAR tomography for large-scale urban mapping," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 5197–5200.
[23] Y. Wang, X. Xiang Zhu, and R. Bamler, "An efficient tomographic inversion approach for urban mapping using meter resolution SAR image stacks," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1250–1254, Jul. 2014.
[24] Y. Shi, X. X. Zhu, W. Yin, and R. Bamler, "A fast and accurate basis pursuit denoising algorithm with application to super-resolving tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6148–6158, Oct. 2018.
[25] A. Budillon, A. C. Johnsy, and G. Schirinzi, "Extension of a fast GLRT algorithm to 5D SAR tomography of urban areas," *Remote Sens.*, vol. 9, no. 8, p. 844, 2017.
[26] J. Otepka, S. Ghuffar, C. Waldhauser, R. Hochreiter, and N. Pfeifer, "Georeferenced point clouds: A survey of features and point cloud management," *ISPRS Int. J. Geo-Inf.*, vol. 2, no. 4, pp. 1038–1065, 2013.
[27] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.
[28] N. Ge, R. Bamler, D. Hong, and X. X. Zhu, "Single-look multimaster SAR tomography: An introduction," *IEEE Trans. Geosci. Remote Sens.*, early access.
[29] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*, 2nd ed. Bellingham, WA, USA: SPIE, 2020.
[30] C.-A. Deledalle, L. Denis, and F. Tupin, "NL-InSAR: Nonlocal interferogram estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1441–1452, Apr. 2011.
[31] X. X. Zhu, G. Baier, M. Lachaise, Y. Shi, F. Adam, and R. Bamler, "Potential and limits of non-local means InSAR filtering for TanDEM-X high-resolution DEM generation," *Remote Sens. Environ.*, vol. 218, pp. 148–161, Dec. 2018.
[32] F. Lombardini and F. Gini, "Model order selection in multi-baseline interferometric radar systems," *EURASIP J. Adv. Signal Process.*, vol. 2005, no. 20, p. 3206, Dec. 2005.
[33] Y. Wang and X. X. Zhu, "Robust estimators for multipass SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 968–980, Feb. 2016.
[34] X. X. Zhu and R. Bamler, "Sparse tomographic SAR reconstruction from mixed TerraSAR-X/TanDEM-X data stacks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 7468–7471.
[35] *Landesamt für Digitalisierung, Breitband und Vermessung Bayern*. Accessed: 2018. [Online]. Available: https://www.ldbv.bayern.de/produkte/3dprodukte/laser.html
[36] M. Schwabisch, "A fast and efficient technique for SAR interferogram geocoding," in *Proc. Sens. Manag. Environ. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 1998, pp. 1100–1102.
[37] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *Int. J. Comput. Vis.*, vol. 13, no. 2, pp. 119–152, Oct. 1994.
[38] I. Sobel, "An isotropic 3×3 image gradient operator," Presented at Stanford AI Project, Stanford, CA, USA, 1968.
[39] Y. Wang, X. X. Zhu, B. Zeisl, and M. Pollefeys, "Fusing meter-resolution 4-D InSAR point clouds and optical images for semantic urban infrastructure monitoring," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 14–26, Jan. 2017.

Yilei Shi (Member, IEEE) received the Diploma (Dipl.-Ing.) degree in mechanical engineering and the Doctorate (Dr.-Ing.) degree in engineering from the Technical University of Munich (TUM), Munich, Germany.

In April and May 2019, he was a Guest Scientist with the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, U.K. He is a Senior Scientist with the Chair of Remote Sensing Technology, Technical University of Munich. His research interests include computational intelligence; fast solver and parallel computing for large-scale problems; advanced methods on SAR and InSAR processing; machine learning and deep learning for a variety of data sources, such as SAR, optical images, and medical images; and PDE-related numerical modeling and computing.

Richard Bamler (Fellow, IEEE) received the Diploma degree in electrical engineering, the Doctorate in Engineering degree, and the Habilitation degree in signal and systems theory from the Technical University of Munich, Munich, Germany, in 1980, 1986, and 1988, respectively.

He worked at the Technical University of Munich from 1981 to 1989 on optical signal processing, holography, wave propagation, and tomography. He joined the German Aerospace Center (DLR), Weßling, Germany, in 1989, where he is the Director of the Remote Sensing Technology Institute. In early 1994, he was a Visiting Scientist with Jet Propulsion Laboratory (JPL), Pasadena, CA, USA, in the preparation of the SIC-C/X-SAR missions, and in 1996, he was a Guest Professor with the University of Innsbruck, Innsbruck, Austria. Since 2003, he has been holding a Full Professorship in remote sensing technology at the Technical University of Munich as a double appointment with his DLR position. His teaching activities include university lectures and courses on signal processing, estimation theory, and SAR. Since he joined DLR, he, his team, and his institute have been working on SAR and optical remote sensing, image analysis and understanding, stereo reconstruction, computer vision, ocean color, passive and active atmospheric sounding, and laboratory spectrometry. They were and are responsible for the development of the operational processors for SIR-C/X-SAR, SRTM, TerraSAR-X, TanDEM-X, Tandem-L, ERS-2/GOME, ENVISAT/SCIAMACHY, MetOp/GOME-2, Sentinel-5P, Sentinel-4, DESIS, and EnMAP. His research interests are in algorithms for optimum information extraction from remote sensing data with emphasis on SAR. This involves new estimation algorithms, such as sparse reconstruction, compressive sensing, and deep learning.

Yuanyuan Wang (Member, IEEE) received the B.Eng. degree (Hons.) in electrical engineering from The Hong Kong Polytechnic University, Hong Kong, in 2008, and the M.Sc. and Dr.-Ing. degrees from the Technical University of Munich (TUM), Munich, Germany, in 2010 and 2015, respectively.

In June and July 2014, he was a Guest Scientist with the Institute of Visual Computing, ETH Zürich, Zürich, Switzerland. He is with Signal Processing in Earth Observation, TUM, as well as with the Department of EO Data Science, German Aerospace Center, Weßling, Germany, where he leads the working group Big SAR Data. His research interests include optimal and robust parameters estimation in multibaseline InSAR techniques, multisensor fusion algorithms of synthetic aperture radar (SAR) and optical data, nonlinear optimization with complex numbers, machine learning in SAR, and high-performance computing for big data.

Dr. Wang was one of the best reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2016.

Xiao Xiang Zhu (Senior Member, IEEE) received the master's (M.Sc.) degree, the Doctor of Engineering (Dr.-Ing.) degree, and the Habilitation degree in signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She is a Professor for Signal Processing in Earth Observation, Technical University of Munich (TUM), and the Head of the Department "EO Data Science," Remote Sensing Technology Institute, German Aerospace Center (DLR), Weßling, Germany. Since 2019, she has been co-coordinating the Munich Data Science Research School. She is also leading the Helmholtz Artificial Intelligence Cooperation Unit (HAICU)—Research Field "Aeronautics, Space and Transport." She was a Guest Scientist or Visiting Professor with the Italian National Research Council (CNR-IREA), Naples, Italy, in 2009, Fudan University, Shanghai, China, in 2014, the University of Tokyo, Tokyo, Japan, in 2015, and the University of California, Los Angeles, Los Angeles, CA, USA, in 2016. Her main research interests are remote sensing, Earth observation, signal processing, machine learning, and data science, with a special application focus on global urban mapping.

Dr. Zhu is a member of young academy (Junge Akademie/Junges Kolleg) at the Berlin-Brandenburg Academy of Sciences and Humanities, the German National Academy of Sciences Leopoldina, and the Bavarian Academy of Sciences and Humanities. She is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.